

```
In [1]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

In [2]: mental_health= pd.read_csv('Mental_Health_in_Pregnancy_During_Covid-19.csv.csv')

In [3]: mental_health.shape

Out[3]: (10772, 16)

In [4]: mental_health.head()

Out[4]:
```

	OSF_ID	Maternal_Age	Household_Income	Maternal_Education	Edinburgh_Postnatal_Depression_Scale	PROMIS_Anxiety	Gestational_Age_At_Birth	Delivery_Date(converted to month and year)
0	1	38.3	\$200,000+	Masters degree	9.0	13.0	39.71	Dec2020
1	2	34.6	\$200,000+	Undergraduate degree	4.0	17.0	NaN	NaN
2	3	34.3	\$100,000 -\$124,999	Undergraduate degree	NaN	NaN	NaN	NaN
3	4	28.8	\$100,000 -\$124,999	Masters degree	9.0	20.0	38.57	Dec2020
4	5	36.5	\$40,000-\$69,999	Undergraduate degree	14.0	20.0	39.86	Oct2020

```
In [5]: mental_health.columns

Out[5]: Index(['OSF_ID', 'Maternal_Age', 'Household_Income', 'Maternal_Education',
              'Edinburgh_Postnatal_Depression_Scale', 'PROMIS_Anxiety',
              'Gestational_Age_At_Birth',
              'Delivery_Date(converted to month and year)', 'Birth_Length',
              'Birth_Weight', 'Delivery_Mode', 'NICU_Stay', 'Language',
              'Threaten_Life', 'Threaten_Baby_Danger', 'Threaten_Baby_Harm'],
              dtype='object')
```

```
In [6]: mental_health.drop(columns=['OSF_ID', 'Language'], inplace=True)

In [7]: mental_health.head()
```

Out[7]:

	Maternal_Age	Household_Income	Maternal_Education	Edinburgh_Postnatal_Depression_Scale	PROMIS_Anxiety	Gestational_Age_At_Birth	Delivery_Date(converted to month and year)	Birth_L
0	38.3	\$200,000+	Masters degree	9.0	13.0	39.71	Dec2020	
1	34.6	\$200,000+	Undergraduate degree	4.0	17.0	NaN	NaN	
2	34.3	\$100,000 - \$124,999	Undergraduate degree	NaN	NaN	NaN	NaN	
3	28.8	\$100,000 - \$124,999	Masters degree	9.0	20.0	38.57	Dec2020	
4	36.5	\$40,000 - \$69,999	Undergraduate degree	14.0	20.0	39.86	Oct2020	

In [8]:

mental_health.isnull().sum()

Out[8]:

Maternal_Age	111
Household_Income	251
Maternal_Education	177
Edinburgh_Postnatal_Depression_Scale	1174
PROMIS_Anxiety	1206
Gestational_Age_At_Birth	4038
Delivery_Date(converted to month and year)	4039
Birth_Length	5292
Birth_Weight	4694
Delivery_Mode	5235
NICU_Stay	5238
Threaten_Life	896
Threaten_Baby_Danger	904
Threaten_Baby_Harm	892
dtype: int64	

In [9]:

mental_health.isna().sum()

```
Out[9]: Maternal_Age                111
Household_Income                  251
Maternal_Education                 177
Edinburgh_Postnatal_Depression_Scale 1174
PROMIS_Anxiety                    1206
Gestational_Age_At_Birth           4038
Delivery_Date(converted to month and year) 4039
Birth_Length                       5292
Birth_Weight                       4694
Delivery_Mode                      5235
NICU_Stay                          5238
Threaten_Life                      896
Threaten_Baby_Danger               904
Threaten_Baby_Harm                 892
dtype: int64
```

```
In [10]: mental_health.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10772 entries, 0 to 10771
Data columns (total 14 columns):
 #   Column                                Non-Null Count  Dtype
---  -
 0   Maternal_Age                         10661 non-null  float64
 1   Household_Income                     10521 non-null  object
 2   Maternal_Education                   10595 non-null  object
 3   Edinburgh_Postnatal_Depression_Scale 9598 non-null   float64
 4   PROMIS_Anxiety                       9566 non-null   float64
 5   Gestational_Age_At_Birth             6734 non-null   float64
 6   Delivery_Date(converted to month and year) 6733 non-null   object
 7   Birth_Length                         5480 non-null   float64
 8   Birth_Weight                         6078 non-null   float64
 9   Delivery_Mode                        5537 non-null   object
10  NICU_Stay                            5534 non-null   object
11  Threaten_Life                        9876 non-null   float64
12  Threaten_Baby_Danger                 9868 non-null   float64
13  Threaten_Baby_Harm                   9880 non-null   float64
dtypes: float64(9), object(5)
memory usage: 1.2+ MB
```

```
In [11]: # Identifying row to drop
rows_to_drop= mental_health[(mental_health['PROMIS_Anxiety'].isnull() & (mental_health['Delivery_Date(converted to month and year)'].isnull())&
                             (mental_health['Threaten_Life'].isna()) & (mental_health['Threaten_Baby_Danger'].isna()) &
                             (mental_health['Threaten_Baby_Harm'].isna()))].index
```

```
In [12]: mental_health.drop(rows_to_drop,inplace= True)
```

```
In [13]: mental_health.shape
```

```
Out[13]: (10292, 14)
```

```
In [14]: (mental_health.isnull().sum()/mental_health.shape[0])*100
```

```
Out[14]: Maternal_Age          0.864749
Household_Income          2.176448
Maternal_Education        1.486592
Edinburgh_Postnatal_Depression_Scale  6.743101
PROMIS_Anxiety            7.054023
Gestational_Age_At_Birth  34.570540
Delivery_Date(converted to month and year) 34.580257
Birth_Length              46.754761
Birth_Weight              40.944423
Delivery_Mode             46.200933
NICU_Stay                 46.230082
Threaten_Life             4.041974
Threaten_Baby_Danger      4.119705
Threaten_Baby_Harm        4.003109
dtype: float64
```

```
In [15]: mental_health.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Index: 10292 entries, 0 to 10771
Data columns (total 14 columns):
#   Column                                Non-Null Count  Dtype
---  -
0   Maternal_Age                          10203 non-null  float64
1   Household_Income                      10068 non-null  object
2   Maternal_Education                    10139 non-null  object
3   Edinburgh_Postnatal_Depression_Scale  9598 non-null   float64
4   PROMIS_Anxiety                       9566 non-null   float64
5   Gestational_Age_At_Birth              6734 non-null   float64
6   Delivery_Date(converted to month and year) 6733 non-null   object
7   Birth_Length                          5480 non-null   float64
8   Birth_Weight                          6078 non-null   float64
9   Delivery_Mode                         5537 non-null   object
10  NICU_Stay                             5534 non-null   object
11  Threaten_Life                         9876 non-null   float64
12  Threaten_Baby_Danger                  9868 non-null   float64
13  Threaten_Baby_Harm                    9880 non-null   float64
dtypes: float64(9), object(5)
memory usage: 1.2+ MB
```

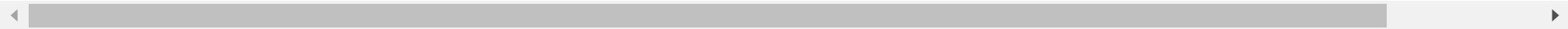
```
In [16]: numerical_df= mental_health.select_dtypes(include=['number'])
```

```
In [17]: numerical_df
```

Out[17]:

	Maternal_Age	Edinburgh_Postnatal_Depression_Scale	PROMIS_Anxiety	Gestational_Age_At_Birth	Birth_Length	Birth_Weight	Threaten_Life	Threaten_Baby_Danger	T
0	38.3	9.0	13.0	39.71	49.20	3431.0	2.0	3.0	
1	34.6	4.0	17.0	NaN	NaN	NaN	2.0	33.0	
3	28.8	9.0	20.0	38.57	41.00	2534.0	53.0	67.0	
4	36.5	14.0	20.0	39.86	53.34	3714.0	23.0	32.0	
5	38.3	3.0	8.0	38.57	NaN	NaN	29.0	36.0	
...	
10766	32.7	15.0	27.0	NaN	NaN	NaN	54.0	64.0	
10767	38.3	10.0	18.0	NaN	NaN	NaN	54.0	62.0	
10769	27.7	4.0	15.0	NaN	NaN	NaN	21.0	93.0	
10770	23.0	12.0	19.0	NaN	NaN	NaN	65.0	94.0	
10771	34.6	18.0	30.0	38.29	NaN	4196.0	84.0	73.0	

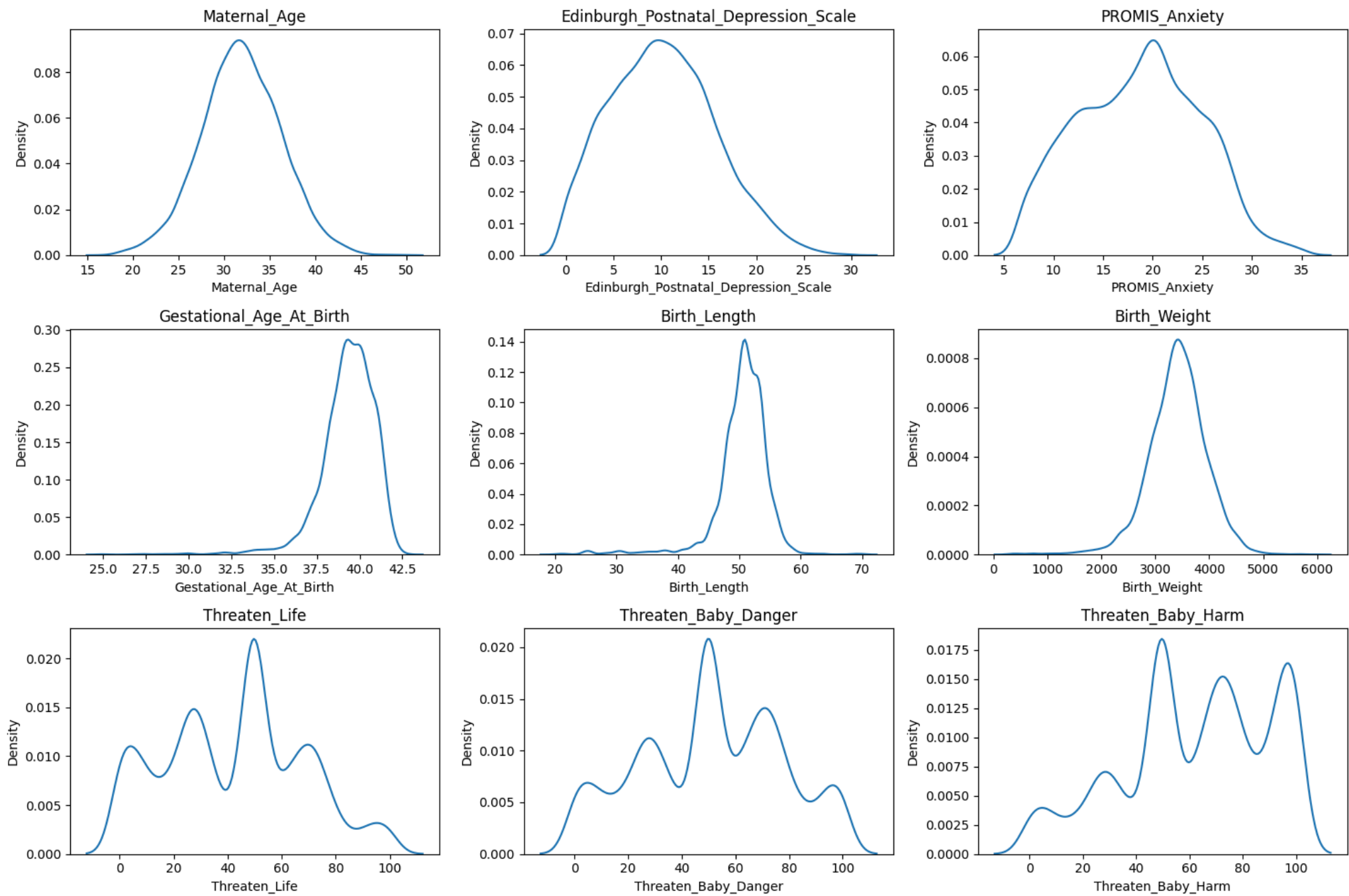
10292 rows × 9 columns



In [18]:

```
fig,ax= plt.subplots(3,3,figsize=(15,10))
ax=ax.flatten()

for i, col in enumerate(numerical_df.columns):
    if i<9:
        sns.kdeplot(data=numerical_df[col],ax=ax[i])
        ax[i].set_title(col)
plt.tight_layout()
plt.show()
```



In [19]: `mental_health.info()`

```

<class 'pandas.core.frame.DataFrame'>
Index: 10292 entries, 0 to 10771
Data columns (total 14 columns):
#   Column                                     Non-Null Count  Dtype
---  -
0   Maternal_Age                             10203 non-null  float64
1   Household_Income                         10068 non-null  object
2   Maternal_Education                       10139 non-null  object
3   Edinburgh_Postnatal_Depression_Scale    9598 non-null   float64
4   PROMIS_Anxiety                          9566 non-null   float64
5   Gestational_Age_At_Birth                 6734 non-null   float64
6   Delivery_Date(converted to month and year) 6733 non-null   object
7   Birth_Length                             5480 non-null   float64
8   Birth_Weight                             6078 non-null   float64
9   Delivery_Mode                             5537 non-null   object
10  NICU_Stay                                5534 non-null   object
11  Threaten_Life                             9876 non-null   float64
12  Threaten_Baby_Danger                      9868 non-null   float64
13  Threaten_Baby_Harm                        9880 non-null   float64
dtypes: float64(9), object(5)
memory usage: 1.2+ MB

```

```

In [20]: print(mental_health['Household_Income'].value_counts())
print('-----')
print(mental_health['Maternal_Education'].value_counts())
print('-----')
print(mental_health['Delivery_Date(converted to month and year)'].value_counts())
print('-----')
print(mental_health['Delivery_Mode'].value_counts())
print('-----')
print(mental_health['NICU_Stay'].value_counts())

```

```
Household_Income
$70,000-$99,999      2006
$100,000 - $124,999  1886
$40,000-$69,999      1348
$125,000- $149,999   1315
$150,000 - $174,999  1115
$200,000+            977
$175,000- $199,999   641
$20,000- $39,999      561
Less than $20, 000    219
Name: count, dtype: int64
```

```
-----
Maternal_Education
Undergraduate degree      3979
College/trade school      2588
Masters degree            1849
High school diploma        835
Doctoral Degree            767
Less than high school diploma  121
Name: count, dtype: int64
```

```
-----
Delivery_Date(converted to month and year)
Aug2020      684
Sep2020      638
Jul2020      630
Oct2020      619
Jun2020      510
Nov2020      473
Jan2021      413
Dec2020      404
Mar2021      385
Apr2021      342
Feb2021      315
May2021      292
Jun2021      281
May2020      214
Jul2021      208
Aug2021      135
Sep2021       86
Oct2021       47
Nov2021       21
Apr2020       12
Sep2022        4
Jun2022        4
Jul2022        4
Dec2021        4
Jan2022        2
Mar2022        2
Aug2022        2
```



```

Feb2022      1
May2022      1
Name: count, dtype: int64
-----
Delivery_Mode
Vaginally          3904
Caesarean-section (c-section)  1633
Name: count, dtype: int64
-----
NICU_Stay
No      4986
Yes     548
Name: count, dtype: int64

```

In [21]: `mental_health.head()`

Out[21]:

	Maternal_Age	Household_Income	Maternal_Education	Edinburgh_Postnatal_Depression_Scale	PROMIS_Anxiety	Gestational_Age_At_Birth	Delivery_Date(converted to month and year)	Birth_L
0	38.3	\$200,000+	Masters degree	9.0	13.0	39.71	Dec2020	
1	34.6	\$200,000+	Undergraduate degree	4.0	17.0	NaN	NaN	
3	28.8	\$100,000 - \$124,999	Masters degree	9.0	20.0	38.57	Dec2020	
4	36.5	\$40,000-\$69,999	Undergraduate degree	14.0	20.0	39.86	Oct2020	
5	38.3	\$150,000 - \$174,999	Undergraduate degree	3.0	8.0	38.57	Jun2020	

In [22]:

```

fig,ax= plt.subplots(1,3,figsize=(15,10))
ax=ax.flatten()

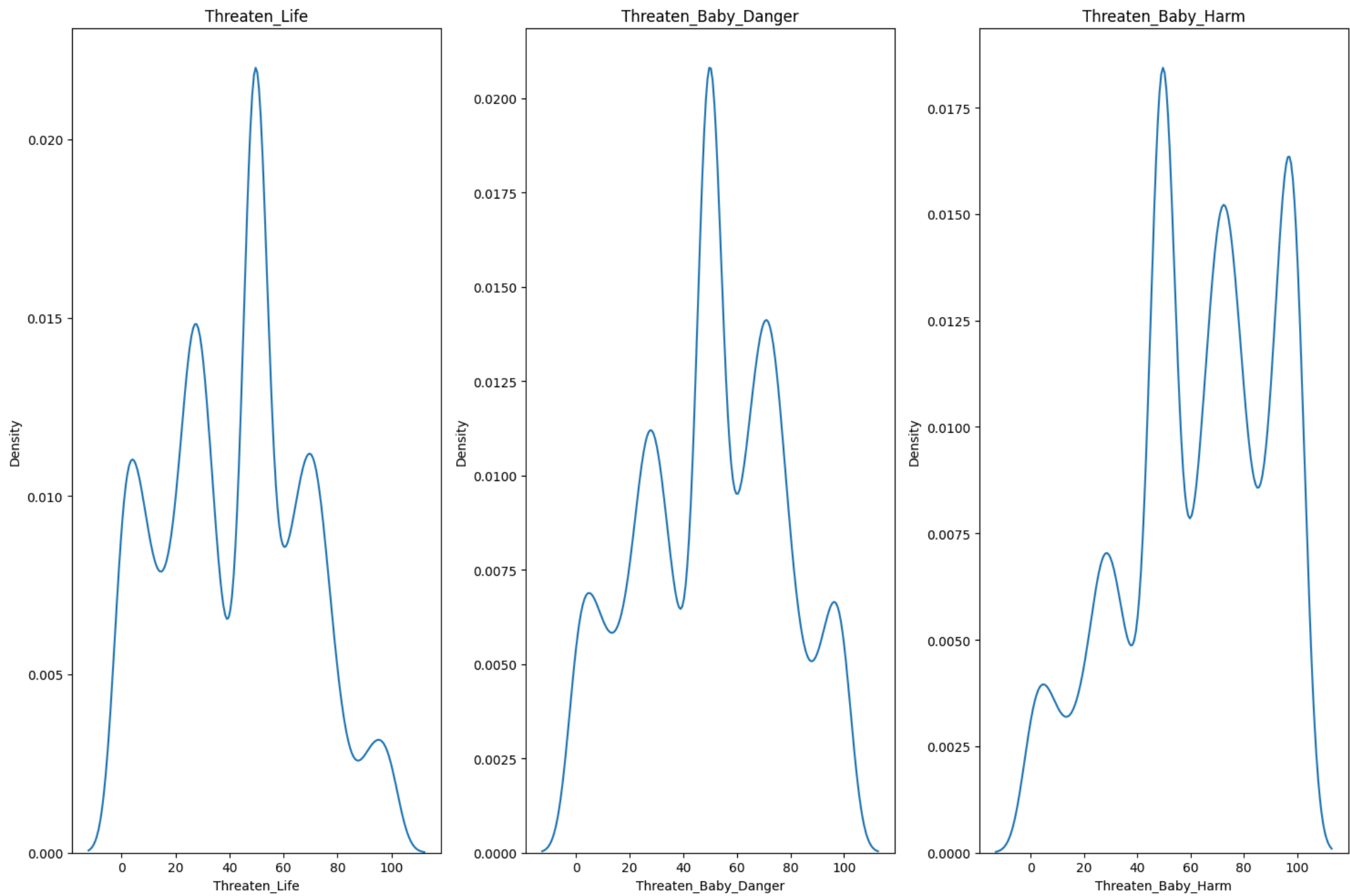
sns.kdeplot(data=mental_health['Threaten_Life'], ax=ax[0])
ax[0].set_title('Threaten_Life')

sns.kdeplot(data=mental_health['Threaten_Baby_Danger'], ax=ax[1])
ax[1].set_title('Threaten_Baby_Danger')

sns.kdeplot(data=mental_health['Threaten_Baby_Harm'], ax=ax[2])
ax[2].set_title('Threaten_Baby_Harm')

plt.tight_layout()
plt.show()

```



Bivariate Analysis

```
In [3]: mental_health.head()
```

Out[3]:

	OSF_ID	Maternal_Age	Household_Income	Maternal_Education	Edinburgh_Postnatal_Depression_Scale	PROMIS_Anxiety	Gestational_Age_At_Birth	Delivery_Date(converted to month and year)
0	1	38.3	\$200,000+	Masters degree	9.0	13.0	39.71	Dec2020
1	2	34.6	\$200,000+	Undergraduate degree	4.0	17.0	NaN	NaN
2	3	34.3	\$100,000 -\$124,999	Undergraduate degree	NaN	NaN	NaN	NaN
3	4	28.8	\$100,000 -\$124,999	Masters degree	9.0	20.0	38.57	Dec2020
4	5	36.5	\$40,000-\$69,999	Undergraduate degree	14.0	20.0	39.86	Oct2020

```
In [4]: # Converted Gestational_Age_At_Birth from weeks to months
mental_health['Gestational_Age_At_Birth']= np.round(mental_health['Gestational_Age_At_Birth']/4.35,2)
```

```
In [5]: mental_health.head()
```

Out[5]:

	OSF_ID	Maternal_Age	Household_Income	Maternal_Education	Edinburgh_Postnatal_Depression_Scale	PROMIS_Anxiety	Gestational_Age_At_Birth	Delivery_Date(converted to month and year)
0	1	38.3	\$200,000+	Masters degree	9.0	13.0	9.13	Dec2020
1	2	34.6	\$200,000+	Undergraduate degree	4.0	17.0	NaN	NaN
2	3	34.3	\$100,000 -\$124,999	Undergraduate degree	NaN	NaN	NaN	NaN
3	4	28.8	\$100,000 -\$124,999	Masters degree	9.0	20.0	8.87	Dec2020
4	5	36.5	\$40,000-\$69,999	Undergraduate degree	14.0	20.0	9.16	Oct2020

```
In [6]: # Lets create column for healthy birth weight

def healthy_weight(row):
    if pd.isnull(row['Birth_Weight']):
        return 'Data Missing'
```

```

elif row['Birth_Weight']>2550 and row['Birth_Weight']<4000:
    return 'Healthy'
else:
    return 'Unhealthy'

# Lets create column for healthy birth length

def healthy_length(row):
    if pd.isnull(row['Birth_Length']):
        return 'Data Missing'
    elif row['Birth_Length']>39 and row['Birth_Length']<61:
        return 'Healthy'
    else:
        return 'Unhealthy'

```

```

In [7]: mental_health['Weight']= mental_health.apply(healthy_weight,axis=1)
        mental_health['Length']= mental_health.apply(healthy_length,axis=1)

```

```

In [8]: mental_health.head()

```

```

Out[8]:
   OSF_ID  Maternal_Age  Household_Income  Maternal_Education  Edinburgh_Postnatal_Depression_Scale  PROMIS_Anxiety  Gestational_Age_At_Birth  Delivery_Date(converted to month and year)

```

	OSF_ID	Maternal_Age	Household_Income	Maternal_Education	Edinburgh_Postnatal_Depression_Scale	PROMIS_Anxiety	Gestational_Age_At_Birth	Delivery_Date(converted to month and year)
0	1	38.3	\$200,000+	Masters degree	9.0	13.0	9.13	Dec2020
1	2	34.6	\$200,000+	Undergraduate degree	4.0	17.0	NaN	NaN
2	3	34.3	\$100,000 -\$124,999	Undergraduate degree	NaN	NaN	NaN	NaN
3	4	28.8	\$100,000 -\$124,999	Masters degree	9.0	20.0	8.87	Dec2020
4	5	36.5	\$40,000-\$69,999	Undergraduate degree	14.0	20.0	9.16	Oct2020



```

In [9]: # Analysis

```

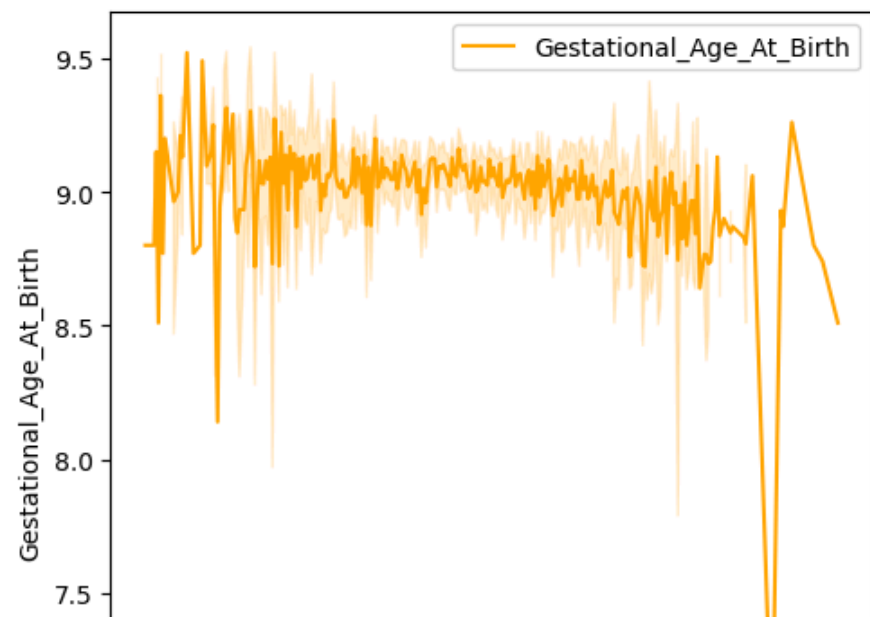
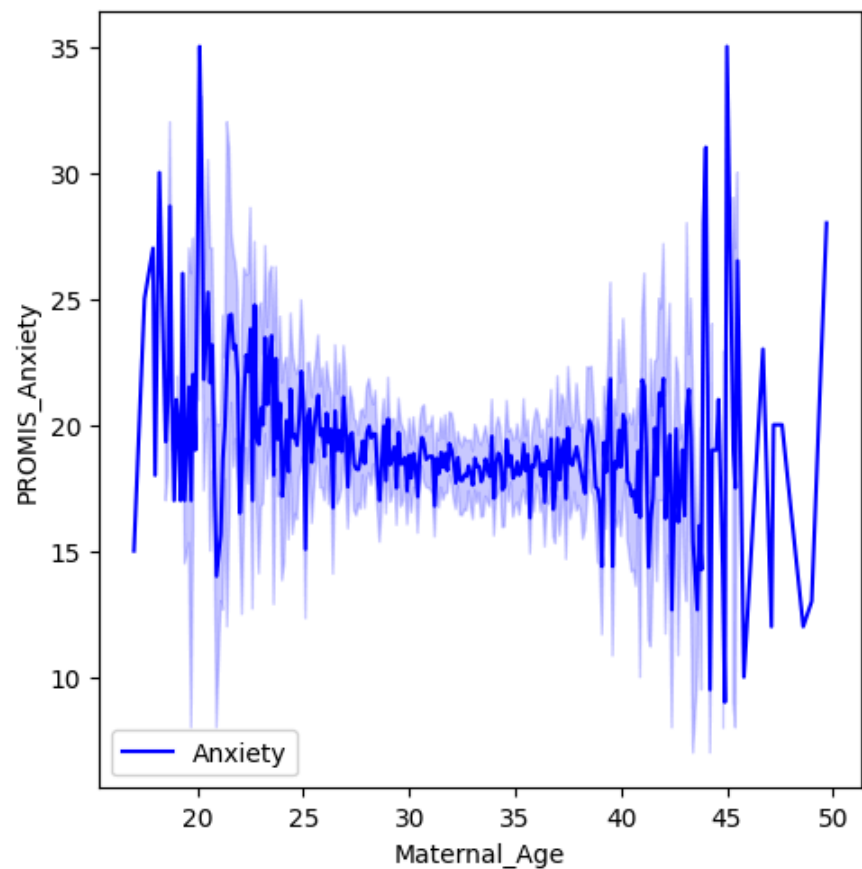
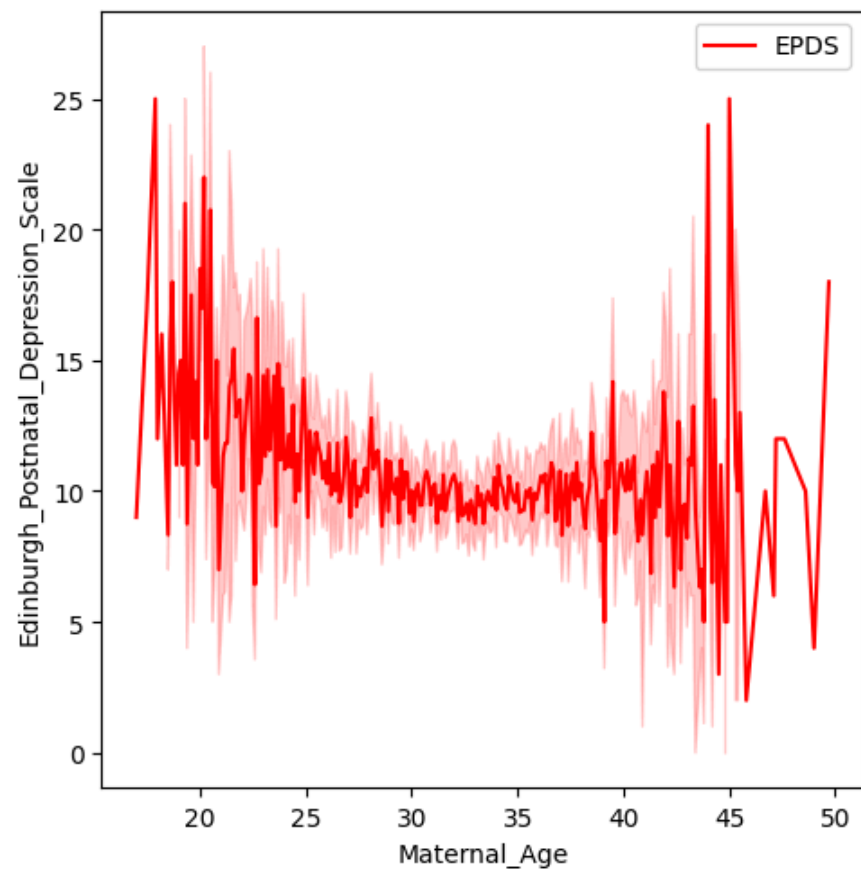
```

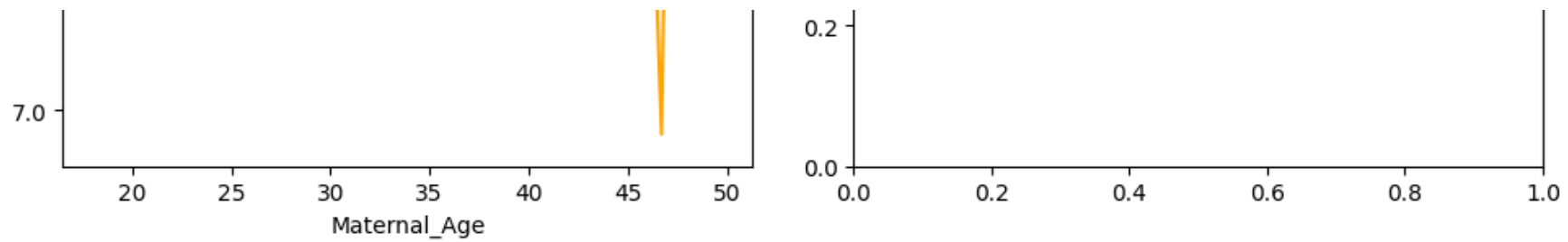
fig, ax=plt.subplots(2,2, figsize=(10,10))
ax= ax.flatten()

sns.lineplot(x='Maternal_Age', y='Edinburgh_Postnatal_Depression_Scale', data=mental_health, ax=ax[0], color='red', label='EPDS')
sns.lineplot(x='Maternal_Age', y='PROMIS_Anxiety', data=mental_health, ax=ax[1], color='blue', label='Anxiety')
sns.lineplot(x='Maternal_Age', y='Gestational_Age_At_Birth', data=mental_health, ax=ax[2], color='orange', label='Gestational_Age_At_Birth')

```

```
plt.tight_layout()  
plt.show()
```

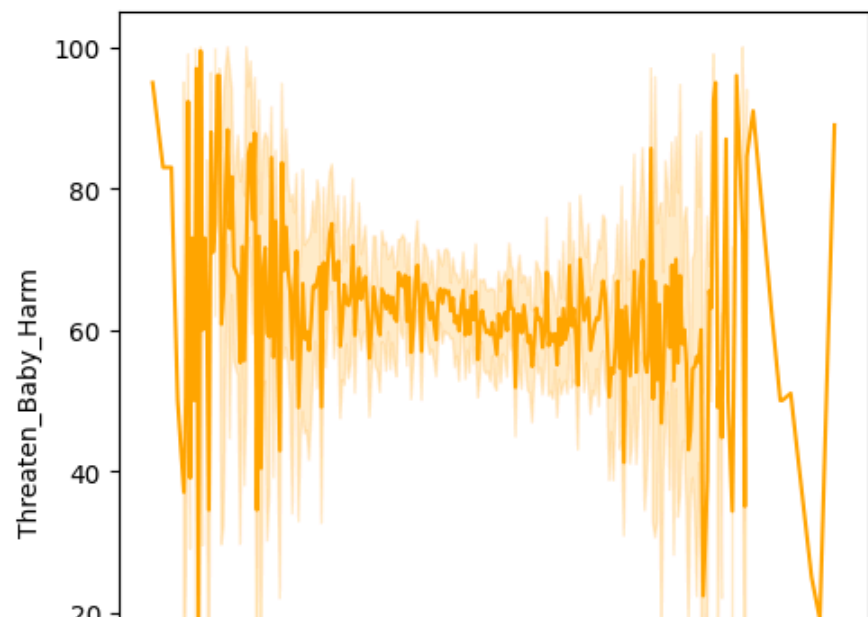
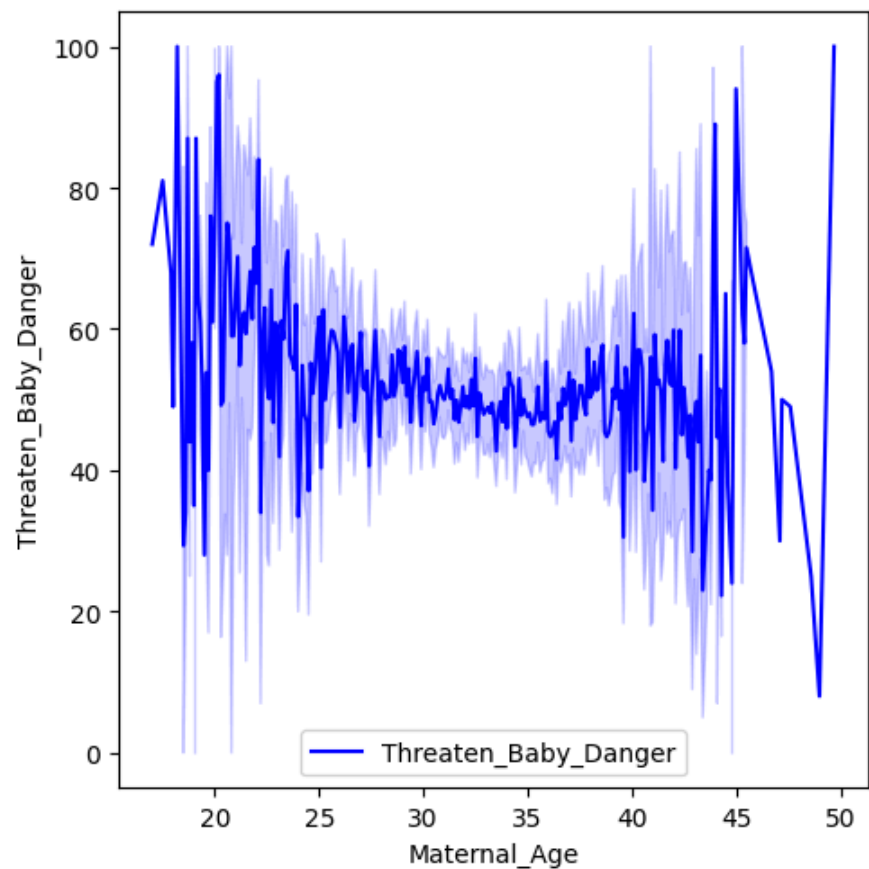
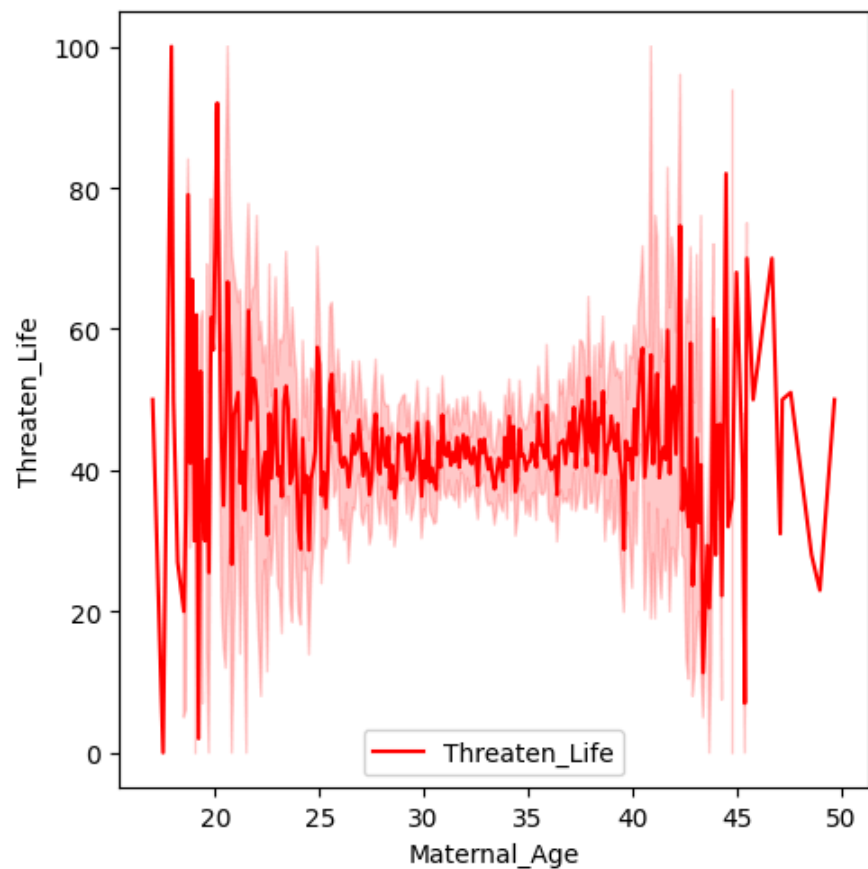


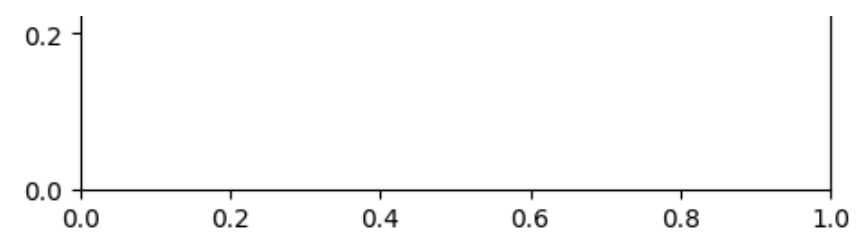
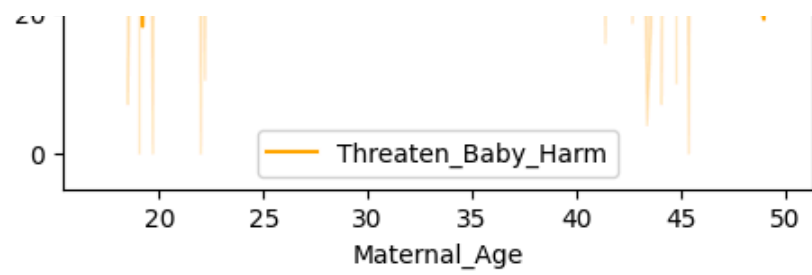


```
In [10]: fig, ax=plt.subplots(2,2, figsize=(10,10))
ax= ax.flatten()

sns.lineplot(x='Maternal_Age', y='Threaten_Life', data=mental_health, ax=ax[0], color='red', label='Threaten_Life')
sns.lineplot(x='Maternal_Age', y='Threaten_Baby_Danger', data=mental_health, ax=ax[1], color='blue', label='Threaten_Baby_Danger')
sns.lineplot(x='Maternal_Age', y='Threaten_Baby_Harm', data=mental_health, ax=ax[2], color='orange', label='Threaten_Baby_Harm')

plt.tight_layout()
plt.show()
```





```
In [11]: new_df= mental_health[mental_health['Delivery_Mode'].notnull()]
```

```
In [12]: new_df
```

Out[12]:

	OSF_ID	Maternal_Age	Household_Income	Maternal_Education	Edinburgh_Postnatal_Depression_Scale	PROMIS_Anxiety	Gestational_Age_At_Birth	Delivery_Date(converted to month and year)	
	0	1	38.3	\$200,000+	Masters degree	9.0	13.0	9.13	Dec 2012
	3	4	28.8	\$100,000 - \$124,999	Masters degree	9.0	20.0	8.87	Dec 2012
	4	5	36.5	\$40,000-\$69,999	Undergraduate degree	14.0	20.0	9.16	Oct 2012
	7	8	NaN	NaN	NaN	NaN	NaN	8.74	May 2012
	8	9	33.1	\$100,000 - \$124,999	College/trade school	1.0	7.0	9.39	Nov 2012

	10755	10756	41.7	\$175,000- \$199,999	Undergraduate degree	19.0	21.0	8.83	Aug 2012
	10756	10757	27.8	\$150,000 - \$174,999	Masters degree	8.0	19.0	8.93	Aug 2012
	10757	10758	36.2	\$150,000 - \$174,999	Undergraduate degree	3.0	9.0	8.87	Jul 2012
	10761	10762	33.2	\$125,000- \$149,999	College/trade school	0.0	8.0	9.56	Oct 2012
	10763	10764	24.0	\$40,000-\$69,999	High school diploma	2.0	13.0	9.52	Dec 2012

5537 rows × 18 columns

```
In [13]: new_df = new_df[['Delivery_Mode', 'Maternal_Age']]
```

```
In [14]: def age_category(row):
    if(row['Maternal_Age']<23):
        return "UnderAge"
    if(row['Maternal_Age']>=23 and row['Maternal_Age']<38):
        return 'Healthy'
    else:
        return 'Old Age'
```

```
In [15]: new_df['Age_Category'] = new_df.apply(age_category,axis=1)
```

```
C:\Users\HP\AppData\Local\Temp\ipykernel_4156\4233722744.py:1: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

```
See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user\_guide/indexing.html#returning-a-view-versus-a-copy
new_df['Age_Category'] = new_df.apply(age_category,axis=1)
```

```
In [16]: new_df
```

```
Out[16]:
```

	Delivery_Mode	Maternal_Age	Age_Category
0	Vaginally	38.3	Old Age
3	Vaginally	28.8	Healthy
4	Caesarean-section (c-section)	36.5	Healthy
7	Caesarean-section (c-section)	NaN	Old Age
8	Vaginally	33.1	Healthy
...
10755	Caesarean-section (c-section)	41.7	Old Age
10756	Caesarean-section (c-section)	27.8	Healthy
10757	Vaginally	36.2	Healthy
10761	Vaginally	33.2	Healthy
10763	Vaginally	24.0	Healthy

5537 rows × 3 columns

```
In [17]: new_df.groupby('Age_Category')['Delivery_Mode'].value_counts().reset_index()
```

```
Out[17]:
```

	Age_Category	Delivery_Mode	count
0	Healthy	Vaginally	3534
1	Healthy	Caesarean-section (c-section)	1381
2	Old Age	Vaginally	327
3	Old Age	Caesarean-section (c-section)	236
4	UnderAge	Vaginally	43
5	UnderAge	Caesarean-section (c-section)	16

```
In [18]: import statsmodels.api as sm
```

```
In [19]: mental_health[['NICU_Stay','Weight','Length']]
```

```
Out[19]:
```

	NICU_Stay	Weight	Length
0	No	Healthy	Healthy
1	NaN	Data Missing	Data Missing
2	NaN	Data Missing	Data Missing
3	No	Unhealthy	Healthy
4	No	Healthy	Healthy
...
10767	NaN	Data Missing	Data Missing
10768	NaN	Data Missing	Data Missing
10769	NaN	Data Missing	Data Missing
10770	NaN	Data Missing	Data Missing
10771	NaN	Unhealthy	Data Missing

10772 rows × 3 columns

```
In [20]: mental_health['NICU_Stay']= mental_health['NICU_Stay'].fillna('Data Missing')
```

```
In [21]: cols= ['NICU_Stay','Weight','Length']
```

```
In [22]: for i in cols:
    #selecting clean data
    df= mental_health[(mental_health[i]!='Data Missing') & (mental_health['Maternal_Age'].notnull())]

    #Binary Conversion
    if i=='NICU_Stay':
        df['binary']= df[i].apply(lambda x : 1 if x=='Yes' else 0)
    else:
        df['binary']= df[i].apply(lambda x : 1 if x=='Unhealthy' else 0)

    # Model
    X= df['Maternal_Age']
    y= df['binary']

    x= sm.add_constant(X)
```

```

logit_model= sm.Logit(y,x)
result= logit_model.fit()

p_value= result.pvalues['Maternal_Age']
alpha= 0.3
if p_value< alpha:
    print(f"There is a significant association between the age of pregnant women and likelihood of {i}")
else:
    print(f"There is no significant association between the age of pregnant women and not likelihood of {i}")

```

Optimization terminated successfully.

Current function value: 0.321824

Iterations 6

There is a significant association between the age of pregnant women and likelihood of NICU_Stay

Optimization terminated successfully.

Current function value: 0.452514

Iterations 6

There is a significant association between the age of pregnant women and likelihood of Weight

Optimization terminated successfully.

Current function value: 0.136026

Iterations 8

There is no significant association between the age of pregnant women and not likelihood of Length

C:\Users\HP\AppData\Local\Temp\ipykernel_4156\537718166.py:7: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.

Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy

```
df['binary']= df[i].apply(lambda x : 1 if x=='Yes' else 0)
```

C:\Users\HP\AppData\Local\Temp\ipykernel_4156\537718166.py:9: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.

Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy

```
df['binary']= df[i].apply(lambda x : 1 if x=='Unhealthy' else 0)
```

C:\Users\HP\AppData\Local\Temp\ipykernel_4156\537718166.py:9: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.

Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy

```
df['binary']= df[i].apply(lambda x : 1 if x=='Unhealthy' else 0)
```

CONCLUSIONS.

- If a lady is in the age group 25-38 while conceiving she will not be under any anxiety attacks , moreover she will also not suffer from any after-pregnancy depression and can take care of herself and her new born more effectively.

- If a female is conceiving beyond the age of 42, this is very risky for the baby as there are very high chances of pre-mature birth of the baby and the pre-mature birth may be in less than 7.5 months of gestation periods.
- Female of the age group 25-42 are less worried about the impact of covid-19 on them and their new born health, damang, they are less worried as compared to females who conceived in less than 25 and females conceiving in after the age 45. its also evident that female after the age of 43 are more worried about the impact of covid-19 on lady and baby , this may be reason for high anxiety and after-pregnency depression in this age group
- After the age of 38, there is a notable 13%[28-41%] increase in the likelihood of women requiring a c-section rather than opting for a natural delivery. this statistic underscores the physiological changes associated with maternal age and their implications for childbirth.
- On testing the relationship between maternal age and [will the new born need intensive medical supervision after birth , healthy or unhealthy weight, healthy or unhealthy height] ,by hypothesis testing , it was found that there is a significant association between the age of pregnant women and the likelihood of NICU_Stay and babe having unhealthy weight, however there is no evidence for significant association between age and height of new born.
- High anxiety , depression , premature birth can be reason for admission of babe to NICU after delivery.

In [23]: *# Household Income*

```
mental_health['Household_Income'].unique()
```

Out[23]: array(['\$200,000+', '\$100,000 -\$124,999', '\$40,000-\$69,999',
 '\$150,000 - \$174,999', nan, '\$70,000-\$99,999',
 '\$125,000- \$149,999', '\$175,000- \$199,999', '\$20,000- \$39,999',
 'Less than \$20, 000'], dtype=object)

In [24]: **import** plotly.express **as** px

#get unique categories of income

```
categories = mental_health['Household_Income'].unique()
```

```
for i in categories:  
    df= mental_health[mental_health['Household_Income']==i]  
    fig= px.histogram(mental_health,x='PROMIS_Anxiety', color='Household_Income',  
                      title='Distribution of PROMIS_Anxiety by household income',  
                      hover_data=mental_health.columns)
```

```
fig.update_layout(  
    xaxis_title='PROMIS_Anxiety',  
    yaxis_title='Count',  
    legend_title='Household Income',  
    barmode='overlay')
```

```
fig.show()
```

```
In [25]: import plotly.express as px

# get unique categories of income

categories = mental_health['Household_Income'].unique()

for i in categories:
    df = mental_health[mental_health['Household_Income']==i]
    fig = px.histogram(mental_health, x='Edinburgh_Postnatal_Depression_Scale', color='Household_Income',
                        title='Distribution of EPDS by household income',
                        hover_data=mental_health.columns)
    fig.update_layout(
        xaxis_title='Edinburgh_Postnatal_Depression_Scale',
        yaxis_title='Count',
        legend_title='Household Income',
        barmode='overlay')
```

```
fig.show()
```

In [26]: `import plotly.express as px`

```
#get unique categories of income
```

```
categories = mental_health['Household_Income'].unique()
```

```
for i in categories:
```

```
    df= mental_health[mental_health['Household_Income']==i]
```

```
    fig= px.histogram(mental_health,x='Threaten_Life', color='Household_Income',
```

```
                      title='Distribution of Threaten_Life by household income',
```

```
                      hover_data=mental_health.columns)
```

```
fig.update_layout(
```

```
    xaxis_title='Threaten_Life',
```

```
    yaxis_title='Count',
```



```
    legend_title='Household Income',  
    barmode='overlay')  
  
fig.show()
```

```
In [27]: mental_health.head()
```

Out[27]:

	OSF_ID	Maternal_Age	Household_Income	Maternal_Education	Edinburgh_Postnatal_Depression_Scale	PROMIS_Anxiety	Gestational_Age_At_Birth	Delivery_Date(converted to month and year)
0	1	38.3	\$200,000+	Masters degree	9.0	13.0	9.13	Dec2020
1	2	34.6	\$200,000+	Undergraduate degree	4.0	17.0	NaN	NaN
2	3	34.3	\$100,000 -\$124,999	Undergraduate degree	NaN	NaN	NaN	NaN
3	4	28.8	\$100,000 -\$124,999	Masters degree	9.0	20.0	8.87	Dec2020
4	5	36.5	\$40,000-\$69,999	Undergraduate degree	14.0	20.0	9.16	Oct2020

In [28]:

```
mental_health.groupby('Household_Income')['Threaten_Life'].mean().reset_index().sort_values('Threaten_Life',ascending=False)
```

Out[28]:

	Household_Income	Threaten_Life
8	Less than \$20, 000	50.771028
4	\$20,000- \$39,999	47.727106
6	\$40,000-\$69,999	45.215491
7	\$70,000-\$99,999	42.388205
0	\$100,000 -\$124,999	41.783813
2	\$150,000 - \$174,999	41.392463
1	\$125,000- \$149,999	40.482463
3	\$175,000- \$199,999	40.062300
5	\$200,000+	39.456635

In [29]:

```
mental_health.groupby('Household_Income')['Threaten_Baby_Danger'].mean().reset_index().sort_values('Threaten_Baby_Danger',ascending=False)
```

Out[29]:

	Household_Income	Threaten_Baby_Danger
--	------------------	----------------------

8	Less than \$20, 000	60.276995
4	\$20,000- \$39,999	58.310786
6	\$40,000-\$69,999	53.740031
7	\$70,000-\$99,999	52.033933
0	\$100,000 -\$124,999	50.288200
1	\$125,000- \$149,999	48.904762
2	\$150,000 - \$174,999	48.456382
3	\$175,000- \$199,999	46.923200
5	\$200,000+	46.885057

```
In [30]: mental_health.groupby('Household_Income')['Threaten_Baby_Harm'].mean().reset_index().sort_values('Threaten_Baby_Harm',ascending=False)
```

Out[30]:

	Household_Income	Threaten_Baby_Harm
--	------------------	--------------------

8	Less than \$20, 000	66.985915
4	\$20,000- \$39,999	66.718464
7	\$70,000-\$99,999	63.633470
6	\$40,000-\$69,999	63.476263
0	\$100,000 -\$124,999	62.386327
1	\$125,000- \$149,999	62.027301
3	\$175,000- \$199,999	60.517572
5	\$200,000+	60.134937
2	\$150,000 - \$174,999	59.983501

CONCLUSIONS

- It is evident from the observations that individuals living below the poverty line (\$70,000) are significantly more susceptible to anxiety and EPDS (Edinburgh Postnatal Depression Scale) scores. Within this income bracket, nearly 70% of the population has been diagnosed with anxiety and depression, exacerbated by the conditions brought on by the COVID-19 pandemic.

- There is a positive correlation between increased income and reduced probability of anxiety and depression. High-income individuals have approximately a 33% chance of being diagnosed with anxiety and depression due to COVID-19, indicating a substantial mental health benefit associated with higher income levels.
- The COVID-19 pandemic has had a profoundly adverse impact on low-income families, particularly those below the survival line. The shortage of income has heightened their worries about the well-being of their children and women. Families below the survival line exhibit a very high threatening score for women, with an average score of 51/100. In contrast, as income levels rise, concerns diminish, likely due to better access to healthcare facilities for higher-income families.
- Unlike the varying threatening scores for women across income levels, the concern for the impact of COVID-19 on newborns is universal. Regardless of income, all categories show significant worry about potential harm and damage to newborns, which may contribute to high anxiety levels in pregnant women during the pandemic.

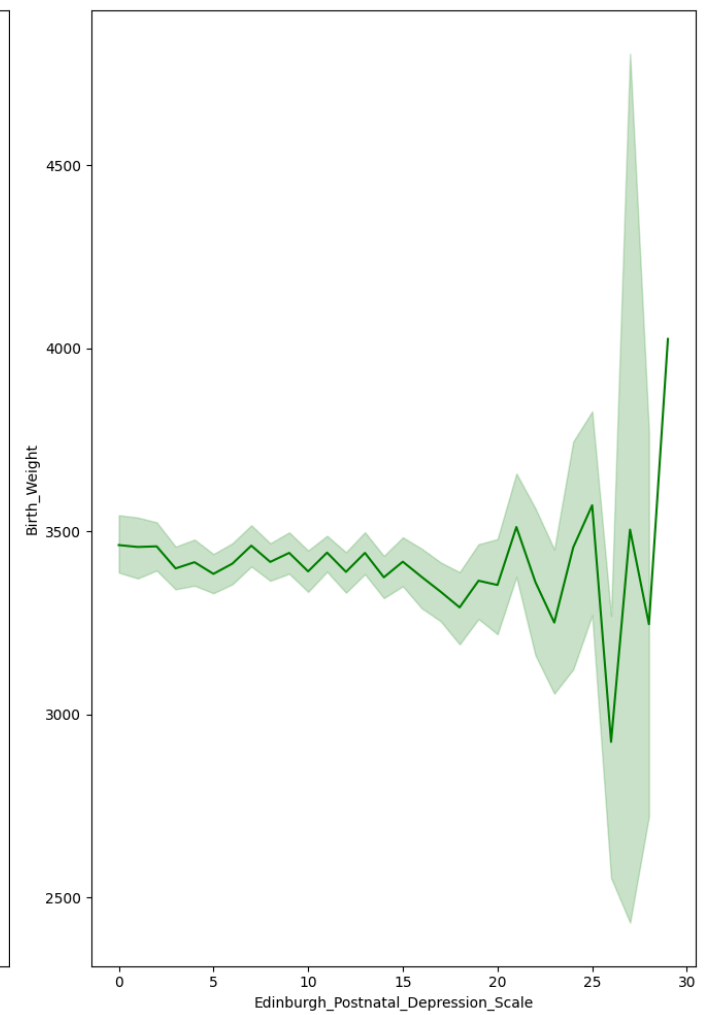
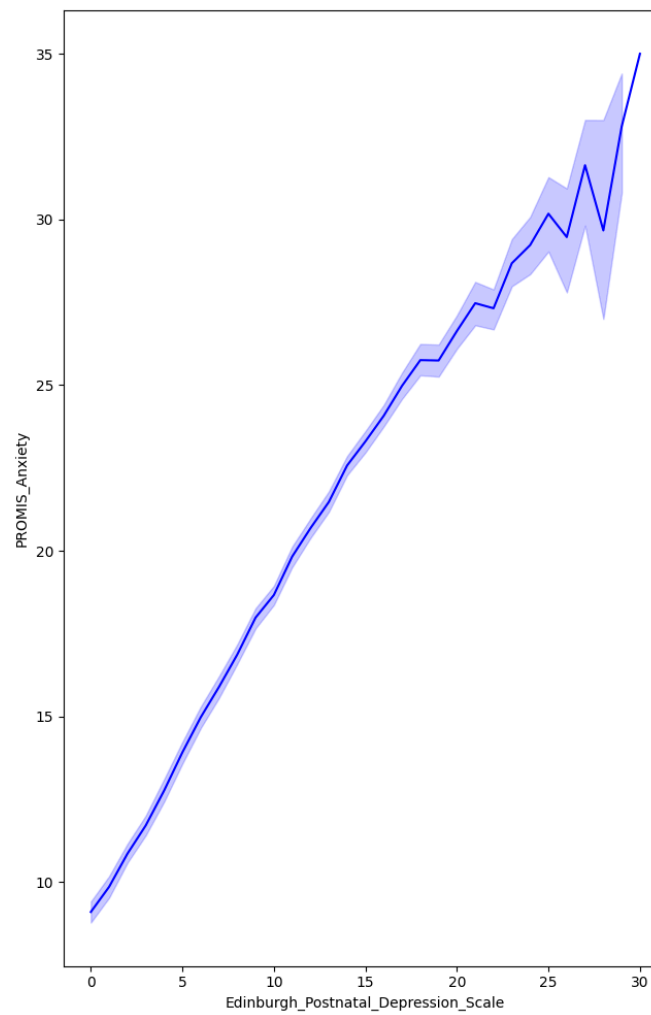
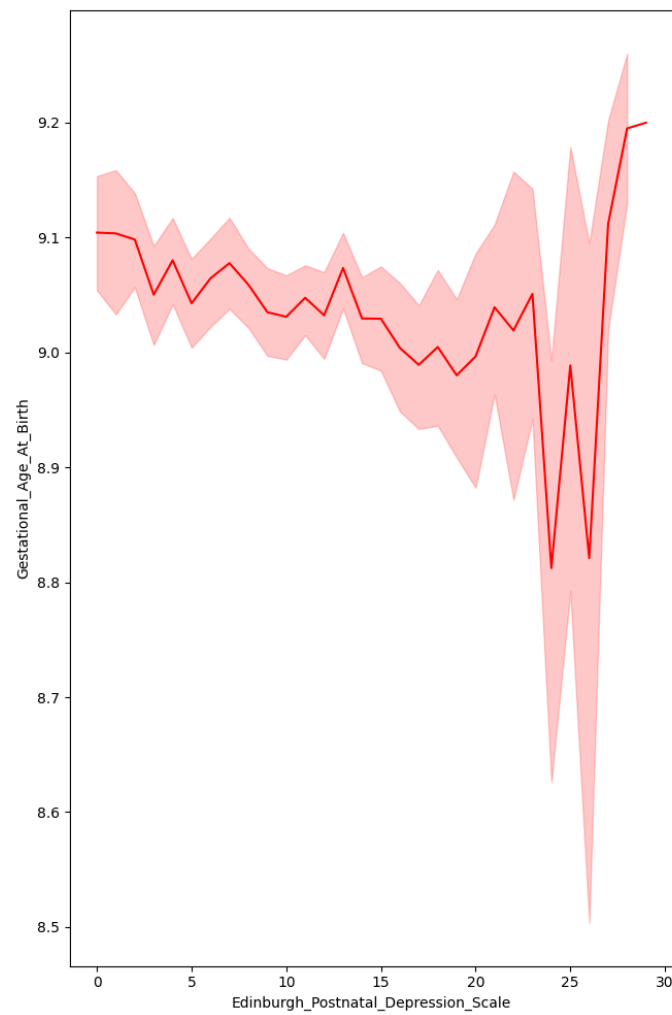
In [62]: `mental_health.head()`

Out[62]:

	Maternal_Age	Household_Income	Maternal_Education	Edinburgh_Postnatal_Depression_Scale	PROMIS_Anxiety	Gestational_Age_At_Birth	Delivery_Date(converted to month and year)	Birth_L
0	38.3	\$200,000+	Masters degree	9.0	13.0	9.13	Dec2020	
1	34.6	\$200,000+	Undergraduate degree	4.0	17.0	NaN	NaN	
3	28.8	\$100,000 - \$124,999	Masters degree	9.0	20.0	8.87	Dec2020	
4	36.5	\$40,000-\$69,999	Undergraduate degree	14.0	20.0	9.16	Oct2020	
5	38.3	\$150,000 - \$174,999	Undergraduate degree	3.0	8.0	8.87	Jun2020	

In [65]:

```
fig, ax= plt.subplots(1,3,figsize=(20,10))
ax= ax.flatten()
sns.lineplot(x='Edinburgh_Postnatal_Depression_Scale', y='Gestational_Age_At_Birth', data=mental_health,
             color='red', ax= ax[0])
sns.lineplot(x='Edinburgh_Postnatal_Depression_Scale', y='PROMIS_Anxiety', data=mental_health,
             color='blue', ax= ax[1])
sns.lineplot(x='Edinburgh_Postnatal_Depression_Scale', y='Birth_Weight', data=mental_health,
             color='green', ax= ax[2])
plt.tight_layout()
plt.show()
```



```
In [66]: df1= mental_health[mental_health['Weight']!='Data Missing']
```

```
In [69]: df1= df1[df1['Delivery_Mode'].notnull()]
```

```
In [71]: # Weight and Deliver Mode
df1.groupby('Weight')['Delivery_Mode'].value_counts().reset_index()
```

Out[71]:

	Weight	Delivery_Mode	count
0	Healthy	Vaginally	3318
1	Healthy	Caesarean-section (c-section)	1286
2	Unhealthy	Vaginally	584
3	Unhealthy	Caesarean-section (c-section)	346

In [72]:

```
In [78]: # NICU Stay and Delivery Mode
df1= mental_health[mental_health['NICU_Stay']!='Data Missing']
df1.groupby('NICU_Stay')['Delivery_Mode'].value_counts().reset_index()
```

Out[78]:

	NICU_Stay	Delivery_Mode	count
0	No	Vaginally	3599
1	No	Caesarean-section (c-section)	1387
2	Yes	Vaginally	304
3	Yes	Caesarean-section (c-section)	244

```
In [76]: # Length and Deliver Mode
df1= mental_health[mental_health['Length']!='Data Missing']
df1= df1[df1['Delivery_Mode'].notnull()]
df1.groupby('Length')['Delivery_Mode'].value_counts().reset_index()
```

Out[76]:

	Length	Delivery_Mode	count
0	Healthy	Vaginally	3763
1	Healthy	Caesarean-section (c-section)	1550
2	Unhealthy	Vaginally	104
3	Unhealthy	Caesarean-section (c-section)	63

```
In [77]: # Length and NICU Saty
df1= mental_health[mental_health['Length']!='Data Missing']
df1= df1[df1['NICU_Stay'].notnull()]
df1.groupby('Length')['NICU_Stay'].value_counts().reset_index()
```

Out[77]:

	Length	NICU_Stay	count
0	Healthy	No	4808
1	Healthy	Yes	505
2	Unhealthy	No	130
3	Unhealthy	Yes	37

CONCLUSIONS

- Observational data indicates a significant correlation between high EPDS scores and the likelihood of pre-mature births. However, it's important to note that these pre-mature births generally occur after 8.5 months of gestation, which falls within the medically accepted range and thus, poses minimal health risks to the newborn.
- Women experiencing high levels of EPDS or anxiety are likely to experience the other condition at similar intensity levels. This linear association suggests that an increase in one condition directly correlates with an increase in the other, highlighting the intertwined nature of these mental health challenges.
- There is a 37% probability that women with babies of unhealthy birth weight will undergo a C-section, which is 10% higher than the likelihood for those with babies of healthy birth weight. This pattern is similarly observed with birth height, indicating a strong link between unhealthy birth metrics and the necessity for C-sections.
- A particularly alarming statistic is that babies with unhealthy birth weight have a 22% chance of requiring a NICU stay, which is 15% higher than for babies with healthy birth weight. This same trend is evident with birth height, underscoring the critical importance of healthy birth metrics in reducing the need for intensive neonatal care.

Multivariate Analysis

In [80]:

mental_health.head()

Out[80]:

	Maternal_Age	Household_Income	Maternal_Education	Edinburgh_Postnatal_Depression_Scale	PROMIS_Anxiety	Gestational_Age_At_Birth	Delivery_Date(converted to month and year)	Birth_L
0	38.3	\$200,000+	Masters degree	9.0	13.0	9.13	Dec2020	
1	34.6	\$200,000+	Undergraduate degree	4.0	17.0	NaN	NaN	
3	28.8	\$100,000 - \$124,999	Masters degree	9.0	20.0	8.87	Dec2020	
4	36.5	\$40,000-\$69,999	Undergraduate degree	14.0	20.0	9.16	Oct2020	
5	38.3	\$150,000 - \$174,999	Undergraduate degree	3.0	8.0	8.87	Jun2020	

◀

▶

In [81]:

Vanurable Group: Thes are the groups that are mostly effected when any pandemic occurs.

In [85]:

mental_health.groupby(
 ['Maternal_Age', 'Household_Income', 'Maternal_Education']
)[['Edinburgh_Postnatal_Depression_Scale']].mean().reset_index().sort_values(
 'Edinburgh_Postnatal_Depression_Scale',ascending=False).head(30)

Out[85]:

	Maternal_Age	Household_Income	Maternal_Education	Edinburgh_Postnatal_Depression_Scale
2644	34.3	\$20,000- \$39,999	College/trade school	29.0
51	20.5	\$40,000-\$69,999	College/trade school	29.0
2037	32.1	\$40,000-\$69,999	High school diploma	28.0
4048	42.2	\$40,000-\$69,999	College/trade school	28.0
992	28.2	Less than \$20, 000	College/trade school	27.0
27	19.6	\$20,000- \$39,999	High school diploma	27.0
42	20.2	Less than \$20, 000	College/trade school	27.0
3274	36.7	\$40,000-\$69,999	Less than high school diploma	26.0
627	26.5	\$20,000- \$39,999	Undergraduate degree	26.0
1714	31.0	\$20,000- \$39,999	High school diploma	26.0
395	25.0	\$70,000-\$99,999	College/trade school	26.0
2169	32.6	\$150,000 - \$174,999	College/trade school	26.0
1038	28.4	\$40,000-\$69,999	High school diploma	26.0
216	23.4	\$40,000-\$69,999	High school diploma	26.0
2001	32.0	\$20,000- \$39,999	Masters degree	25.0
3803	39.5	\$20,000- \$39,999	College/trade school	25.0
1685	30.9	\$20,000- \$39,999	Undergraduate degree	25.0
4135	45.0	\$70,000-\$99,999	College/trade school	25.0
968	28.1	Less than \$20, 000	College/trade school	25.0
942	28.0	\$40,000-\$69,999	Less than high school diploma	25.0
2	17.9	Less than \$20, 000	Less than high school diploma	25.0
1273	29.3	\$70,000-\$99,999	High school diploma	24.5
296	24.2	Less than \$20, 000	Undergraduate degree	24.0
689	26.8	\$40,000-\$69,999	Less than high school diploma	24.0
186	23.0	\$70,000-\$99,999	College/trade school	24.0
4119	44.0	\$40,000-\$69,999	Undergraduate degree	24.0

	Maternal_Age	Household_Income	Maternal_Education	Edinburgh_Postnatal_Depression_Scale
3860	40.0	\$150,000 - \$174,999	Doctoral Degree	24.0
3904	40.3	\$40,000-\$69,999	Masters degree	24.0
10	18.6	\$40,000-\$69,999	Less than high school diploma	24.0
3455	37.5	\$70,000-\$99,999	High school diploma	24.0

In [88]: *# Lets see the probability of women with low EPDS having high Anxiety*

```
a= len(mental_health[mental_health['Edinburgh_Postnatal_Depression_Scale']<12])
b= len(mental_health[(mental_health['Edinburgh_Postnatal_Depression_Scale']<12) & (mental_health['PROMIS_Anxiety']>20)])
```

In [89]: `print(f"Probability of women having low EPDS and High Anxiety is {(b/a)*100}")`

Probability of women having low EPDS and High Anxiety is 14.884768670940913

In [91]: *# Lets see the probability of women with high EPDS having high Anxiety*

```
a= len(mental_health[mental_health['Edinburgh_Postnatal_Depression_Scale']>12])
b= len(mental_health[(mental_health['Edinburgh_Postnatal_Depression_Scale']>12) & (mental_health['PROMIS_Anxiety']>20)])
print(f"Probability of women having low EPDS and High Anxiety is {(b/a)*100}")
```

Probability of women having low EPDS and High Anxiety is 81.31833801936895

In [92]: `mental_health.head()`

Out[92]:

	Maternal_Age	Household_Income	Maternal_Education	Edinburgh_Postnatal_Depression_Scale	PROMIS_Anxiety	Gestational_Age_At_Birth	Delivery_Date(converted to month and year)	Birth_L
--	--------------	------------------	--------------------	--------------------------------------	----------------	--------------------------	--	---------

0	38.3	\$200,000+	Masters degree	9.0	13.0	9.13	Dec2020	
1	34.6	\$200,000+	Undergraduate degree	4.0	17.0	NaN	NaN	
3	28.8	\$100,000 - \$124,999	Masters degree	9.0	20.0	8.87	Dec2020	
4	36.5	\$40,000-\$69,999	Undergraduate degree	14.0	20.0	9.16	Oct2020	
5	38.3	\$150,000 - \$174,999	Undergraduate degree	3.0	8.0	8.87	Jun2020	



In [93]: `df= mental_health.copy()`

```
In [ ]: df= df[(df['Threaten_Life']!=' ')&(df['Threaten_Baby_Harm']!=' ')&(df['Threaten_Baby_Danger']!=' ')]
```

```
In [97]: df.groupby(  
    ['Maternal_Age','Household_Income','Maternal_Education']  
) .agg({'Threaten_Life':'mean','Threaten_Baby_Harm':'mean','Threaten_Baby_Danger':'mean'}).reset_index().sort_values(  
    by= ['Threaten_Life','Threaten_Baby_Harm','Threaten_Baby_Danger'], ascending=False  
) .head(40)
```

Out[97]:

	Maternal_Age	Household_Income	Maternal_Education	Threaten_Life	Threaten_Baby_Harm	Threaten_Baby_Danger
54	20.6	\$20,000- \$39,999	High school diploma	100.0	100.0	100.0
529	25.9	\$20,000- \$39,999	College/trade school	100.0	100.0	100.0
882	27.8	\$100,000 -\$124,999	Masters degree	100.0	100.0	100.0
905	27.9	\$100,000 -\$124,999	College/trade school	100.0	100.0	100.0
940	28.0	\$40,000-\$69,999	College/trade school	100.0	100.0	100.0
977	28.2	\$150,000 - \$174,999	College/trade school	100.0	100.0	100.0
1475	30.1	Less than \$20, 000	College/trade school	100.0	100.0	100.0
1573	30.5	\$20,000- \$39,999	High school diploma	100.0	100.0	100.0
2268	32.9	\$40,000-\$69,999	High school diploma	100.0	100.0	100.0
2653	34.3	\$40,000-\$69,999	High school diploma	100.0	100.0	100.0
2864	35.1	\$20,000- \$39,999	Less than high school diploma	100.0	100.0	100.0
3522	37.8	Less than \$20, 000	Masters degree	100.0	100.0	100.0
4048	42.2	\$40,000-\$69,999	College/trade school	100.0	100.0	100.0
4113	43.8	\$20,000- \$39,999	College/trade school	100.0	100.0	98.0
4069	42.6	Less than \$20, 000	Undergraduate degree	100.0	100.0	96.0
1226	29.1	Less than \$20, 000	Less than high school diploma	100.0	100.0	95.0
1236	29.2	\$175,000- \$199,999	Doctoral Degree	100.0	100.0	86.0
4073	42.7	\$125,000- \$149,999	Undergraduate degree	100.0	100.0	85.0
2726	34.6	\$20,000- \$39,999	High school diploma	100.0	100.0	72.0
881	27.8	\$100,000 -\$124,999	High school diploma	100.0	98.0	100.0
2001	32.0	\$20,000- \$39,999	Masters degree	100.0	97.0	93.0
3932	40.7	\$125,000- \$149,999	Doctoral Degree	100.0	96.0	95.0
3216	36.5	\$100,000 -\$124,999	Masters degree	100.0	96.0	NaN
3226	36.5	\$20,000- \$39,999	College/trade school	100.0	95.0	99.0
2276	33.0	\$100,000 -\$124,999	Doctoral Degree	100.0	95.0	96.0
3945	40.8	\$175,000- \$199,999	College/trade school	100.0	94.0	89.0

	Maternal_Age	Household_Income	Maternal_Education	Threaten_Life	Threaten_Baby_Harm	Threaten_Baby_Danger
2814	34.9	\$40,000-\$69,999	College/trade school	100.0	93.0	94.0
2872	35.1	\$70,000-\$99,999	Masters degree	100.0	92.0	96.0
3665	38.6	\$150,000 - \$174,999	High school diploma	100.0	88.0	92.0
320	24.4	Less than \$20, 000	College/trade school	100.0	88.0	78.0
2	17.9	Less than \$20, 000	Less than high school diploma	100.0	83.0	68.0
216	23.4	\$40,000-\$69,999	High school diploma	100.0	79.0	100.0
2346	33.2	\$175,000- \$199,999	Undergraduate degree	100.0	70.0	60.0
1429	29.9	Less than \$20, 000	High school diploma	100.0	NaN	NaN
473	25.5	\$70,000-\$99,999	Undergraduate degree	99.0	100.0	98.0
485	25.6	\$40,000-\$69,999	Undergraduate degree	99.0	96.0	97.0
770	27.2	\$70,000-\$99,999	High school diploma	99.0	95.0	99.0
2120	32.4	\$20,000- \$39,999	Masters degree	99.0	94.0	97.0
189	23.1	\$20,000- \$39,999	College/trade school	99.0	94.0	71.0
372	24.8	Less than \$20, 000	College/trade school	99.0	11.0	95.0

In []:

CONCLUSIONS

- The most vulnerable group to high scores on the Edinburgh Postnatal Depression Scale (EPDS) consists of women of all age groups with low household income and low education levels. However, there are notable instances where women from high-income and highly educated backgrounds are also susceptible to high EPDS scores. This vulnerability predominantly affects women aged 40 and above, suggesting that advanced maternal age can be a significant risk factor for depression during pregnancy and postpartum.
- There is a 14% likelihood that women with low EPDS (depression) scores will experience high anxiety. Conversely, women with high EPDS scores have an 81% chance of also experiencing high anxiety, particularly exacerbated by the stress and uncertainties brought on by the COVID-19 pandemic. This strong correlation underscores the compounded mental health challenges faced by this group during such crises.
- Families with low income, regardless of age and education levels, exhibit heightened concerns about the health and safety of the mother and newborn due to COVID-19. This widespread anxiety reflects the profound impact of socioeconomic status on perceived vulnerability and health security during the pandemic.

- Individuals with a master's degree but low income, and those with high income but lower education levels, share similar levels of concern about the impact of COVID-19 on the mother and newborn. This finding highlights that both education and income are equally significant factors in determining the level of worry and perceived risk related to the pandemic's effects on maternal and infant health.

In []: